

## HORIZON RUN 4 SIMULATION: COUPLED EVOLUTION OF GALAXIES AND LARGE-SCALE STRUCTURES OF THE UNIVERSE

JUHAN KIM<sup>a</sup>, CHANGBOM PARK<sup>b</sup>, BENJAMIN L'HUIILLIER<sup>b,\*</sup>, AND SUNGWOOK E. HONG<sup>b</sup>

<sup>a</sup> Center for Advanced Computation, Korea Institute for Advanced Study, 85 Hoegiro, Dongdaemun-gu, Seoul 130-722, Korea

<sup>b</sup> School of Physics, Korea Institute for Advanced Study, 85 Hoegiro, Dongdaemun-gu, Seoul 130-722, Korea

Received –, 2015; accepted –, 2015

**Abstract:** The Horizon Run 4 is a cosmological  $N$ -body simulation designed for the study of coupled evolution between galaxies and large-scale structures of the Universe, and for the test of galaxy formation models. Using  $6300^3$  gravitating particles in a cubic box of  $L_{\text{box}} = 3150 h^{-1}\text{Mpc}$ , we build a dense forest of halo merger trees to trace the halo merger history with a halo mass resolution scale down to  $M_s = 2.7 \times 10^{11} h^{-1}\text{M}_\odot$ . We build a set of particle and halo data, which can serve as testbeds for comparison of cosmological models and gravitational theories with observations. We find that the FoF halo mass function shows a substantial deviation from the universal form with tangible redshift evolution of amplitude and shape. At higher redshifts, the amplitude of the mass function is lower, and the functional form is shifted toward larger values of  $\ln(1/\sigma)$ . We also find that the baryonic acoustic oscillation feature in the two-point correlation function of mock galaxies becomes broader with a peak position moving to smaller scales and the peak amplitude decreasing for increasing directional cosine  $\mu$  compared to the linear predictions. From the halo merger trees built from halo data at 75 redshifts, we measure the half-mass epoch of halos and find that less massive halos tend to reach half of their current mass at higher redshifts. Simulation outputs including snapshot data, past lightcone space data, and halo merger data are available at <http://sdss.kias.re.kr/astro/Horizon-Run4/>.

**Key words:** cosmology: large-scale structure of universe — galaxies: evolution — methods: numerical

### 1. INTRODUCTION

Over the last decade, a series of cosmological  $N$ -body simulations called the Horizon Run (HRs) simulations have served as a testbed for cosmological models through comparisons with the observed large-scale distribution of galaxies. The first Horizon Run (HR1) was performed in 2008 and published in 2009 (Kim et al. 2009). The simulation box size was  $L_{\text{box}} = 6592 h^{-1}\text{Mpc}$ , and the number of evolved particles was  $N_p = 4120^3$ . The initial power spectrum was calculated by a fitting function from Eisenstein & Hu (1998), adopting a standard  $\Lambda$  cold dark matter ( $\Lambda\text{CDM}$ ) cosmology in a concordance with WMAP 5-year observations (Dunkley et al. 2009). It produced eight non-overlapping all-sky lightcone data of halos and subhalos up to  $z = 0.6$ . We studied the non-linear gravitational effects on the baryonic acoustic oscillation (BAO) peak by measuring the changes in the peak position and amplitude. In 2011, we performed even bigger simulations called Horizon Run 2 and 3 (HR2 and HR3, respectively; Kim et al. 2011). By adopting the same cosmological model as in HR1, the initial power spectra of HR2 and HR3 were generated from the CAMB package (Lewis et al. 2000). The simulated galaxy distributions have been extensively exploited to measure both the expected distribution of the largest structures for testing

cosmic homogeneity (Park et al. 2012) and the cosmic topology for constraining the non-linear gravitational effect on the halo density field (Choi et al. 2013; Kim et al. 2014; Speare et al. 2015).

All the previous HRs have mean particle separations larger than  $1 h^{-1}\text{Mpc}$ , which has been sufficient for many cosmological tests. With much success in quantifying the non-linear gravitational effects on large-scale structures, recently we extended our research focus to galaxy formation studies. To model galaxies in simulations, we employed the subhalo-galaxy one-to-one correspondence model and abundance matching between subhalo mass function and the observed galaxy luminosity or stellar mass function (Kim et al. 2008). Most characteristics of observed galaxy distributions (in terms of luminosity functions and one-point density distributions) are well reproduced by the model while observed abundances of galaxy clusters are not properly recovered from subhalos. The underpopulation of simulated galaxy clusters may come from the inefficient subhalo findings in cluster regions (Muldrew et al. 2011; for the various subhalo finding comparisons see Onions et al. 2012) or from the spatial decoupling between subhalos and galaxies. The latter may survive the tidal disruption longer due to more compact sizes through baryonic dissipation (Weinberg et al. 2008).

In the  $\Lambda\text{CDM}$  cosmology, dark matter halos form hierarchically through the merger of smaller structures. These merger events can trigger star formation and

CORRESPONDING AUTHOR:  
 lhuillier@kias.re.kr

B. L'Huillier; e-mail:

drive galaxy formation and evolution (Kauffmann et al. 2004; Blanton & Berlind 2007). The merger history of galaxies has extensively been studied in semi-analytic models (SAMs; Cole et al. 1994; Kauffmann et al. 1997; De Lucia et al. 2004; Baugh 2006; Lee et al. 2014) for the last two decades. In SAMs, the gas heating and cooling rate are tabulated, and the resulting star formation and supernovae feedback effects are implemented with some parametric prescriptions. Those parameters are fine-tuned to reproduce the correlation functions and/or luminosity functions of observed galaxies. Even though SAMs have achieved a great success in reproducing some observables, they require the introduction of a large number of parameters that are not necessarily physical.

Another well-known empirical galaxy model, the halo occupation distribution (HOD; Berlind & Weinberg 2002; Zheng et al. 2005, 2009) modeling, has been adopted to match the inner part of the galaxy correlation, which is attributed to satellite pairs inside a virialized halo. To distribute satellite galaxies in a halo, they empirically measure the probability number distribution of satellites from observations. The HOD is simpler than SAMs, and widely used for the comparison between observed galaxies and simulated halos.

The galaxy-subhalo correspondence (or the abundance matching; Kim et al. 2008; Trujillo-Gomez et al. 2011; Rodriguez-Puebla et al. 2013; Reddick et al. 2013; Klypin et al. 2015) model is positioned between the two aforementioned models. It is much simpler than SAMs but based on more physical processes than the HOD. It originally models the satellite galaxy distribution from subhalo catalogs. Satellite galaxies in a galaxy cluster originally formed *in situ* isolated and merged into the cluster afterward. While falling into the potential well of the cluster, they experience a drag force by dynamical friction (Zhao 2004), and they inevitably show spiraling inward orbital motions. After a certain time, they finally merge into the central galaxy.

Although it seems reasonable to assume the presence of a satellite galaxy inside a subhalo as long as there is no galaxy-halo decoupling, it has been noted that some satellite galaxies may not have a host subhalo (Gao et al. 2004; Guo et al. 2011; Guo & White 2014; Wang et al. 2014). This could be tested by extensive hydrodynamical simulations to investigate the segregation between satellite galaxies and their dark matter host (Weinberg et al. 2008). However, hydrodynamical simulations are expensive to run and still require much effort to reduce ambiguities in astrophysical processes and numerical artifacts. On the other hand, Hong et al. (2015) recently showed that if most bound particles are used instead of subhalos in the modeling, it is possible to identify such satellites without hosting subhalo.

Therefore, we performed a new simulation in our series, the Horizon Run 4 (HR4). This simulation, with improved spatial and mass resolutions with respect to the previous runs, retains a large number of particles. It is well-suited to study galaxy formation by producing merger trees.

The outline of the paper is as follows. In Section 2 and 3, we describe the simulation specifics and outputs of HR4, respectively. Mass function, shape and spin of virialized halos are dealt with in Section 4. The analysis of two-point correlation functions and mass accretion history are given in Section 5 and 6, respectively. Summary and discussions are following in Section 7.

## 2. GOTPM CODE AND SIMULATION

### 2.1. Initial Conditions & Parallelism

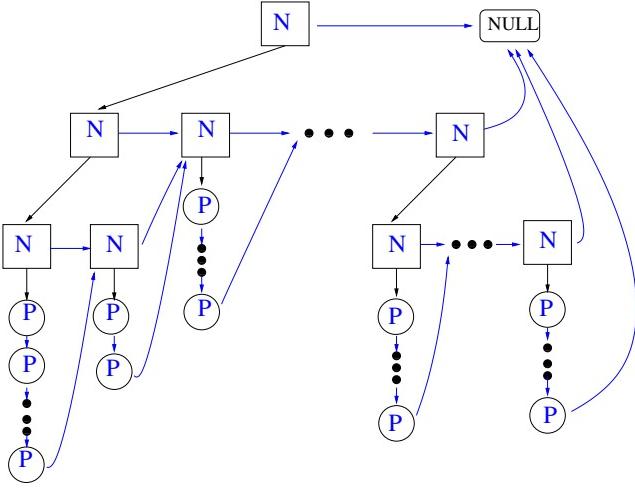
The simulation was run with an improved version of the GOTPM code (Dubinski et al. 2004). The input power spectrum is calculated by the CAMB package, and the initial positions and velocities of the particles are calculated by applying the second-order Lagrangian perturbation theory (2LPT) method proposed by Jenkins (2010). The gravitational force is evaluated through splitting the Newtonian force law into long- and short-range forces (for the Newtonian and Relativistic relations, see Rigopoulos & Valkenburg 2015; Hwang et al. 2012). The long-range forces are calculated from the Poisson equation in Fourier space for the density mesh built by the Particle-Mesh (PM) method. The short-range forces are measured with the Tree method.

We parallelized the GOTPM code implementing MPI and OpenMP with a one-dimensional domain decomposition ( $z$ -directional slabs). We adopt a dynamic domain decomposition, which sets the number of particle in each domain to be equal within one percent. Accordingly, the slab width changes during the simulation run. By using a dynamic domain decomposition, one can easily identify the neighborhoods of a domain and establish communications between them. On the other hand, slab domains usually have greater surface-to-volume ratios than ordinary cubic domains (e.g., the orthogonal recursive bisection, Dubinski 1996), and so it has large communication size between domains.

### 2.2. Non-recursive Oct-Sibling Tree

We have employed a non-recursive oct-sibling tree (OST) for the tree-force update. The OST is a structured tree of particles and nodes with mutual connections established by sibling and daughter pointers. Each particle has one sibling pointer, and each node has two pointers: one for its daughter and the other for the next sibling.

First, we create a top-most node encompassing all particles. We define it as the zero tree level, and its sibling pointer is directed to the null value (Fig. 1). From the top-most node, we recursively divide each node into eight equal-sized cubic subnodes by increasing the tree level by one. If a sibling subnode contains more particles than a predefined number, we divide the node further by increasing the tree level by one. The daughter pointer of the node is directed to the first sibling subnode, and the other subnodes are linked by their sibling pointers. If the node does not have any subnode, we make a chain of particles linked by their sibling pointers, and we set the start and end of the chain connected



**Figure 1.** Example of the Oct-Sibling Tree structure. Boxes and circles represent nodes and particles, respectively. The black and blue arrows are daughter and sibling pointers, respectively. Each node has a daughter and sibling pointers while each particle has only a sibling pointer.

to the previous and next sibling nodes (or possibly particles), respectively. The last sibling at each local tree level is set to have its sibling pointer directed to the mother’s next sibling if it exists. If not, we also recursively climb the local tree until we find the next sibling of the current tree line.

The advantage of the OST over the traditional oct-tree is a smaller number of pointers it employs and the needlessness of a recursive tree walk, which requires additional costs for such a stacking process to temporarily store information of the current recursive depth. The algorithm 2.1 is a pseudocode for the non-recursive tree walk with the OST. The tree walk is taken until the running pointer,  $p$ , encounters the null value. During the tree walks the opening of a node is determined by the **Open** function. The tree-force update is done either by **GroupForce** or **ParticleForce** depending on the data type addressed by the pointer ( $p \rightarrow \text{type}$ ). These three functions play a pivotal role in tree walks. **Open** decides whether to go further into one deeper level (opening the node and going down to its daughter) or jump to the next sibling under the opening condition that  $\theta > \theta_c$ , with  $\theta_c$  the predefined opening threshold. The **GroupForce** function calculates the gravitational force from the group of particles using the multipole expansion while **ParticleForce** calculates the gravitational force from the particle,  $p$ . Thanks to its cost efficiency, this kind of pseudocode is widely applied to our analysis tools such as a percolation method (Friend-of-Friend halo finding), peak findings in a Spherical Overdensity halo identification, and the two-point correlation measurement.

### 2.3. Position Accuracy in GOTPM

One of the key factors to determine the resolution of Lagrangian codes is the spatial accuracy or, more specifically, the number of significant digits involved in the

### Algorithm 2.1: GOTPMTREEWALK( $p$ )

```

while  $p \neq \text{NULL}$ 
  if ( $p \rightarrow \text{type}$ ) = NODE
    if Open( $p$ ) = YES
      then  $\{ p := p \rightarrow \text{daughter}$ 
      else  $\{ \text{call } \text{GroupForce}(p)$ 
    else  $\{ \text{call } \text{ParticleForce}(p)$ 
  comment:  $p$  is a running pointer.

```

particle position. Usually, a single-precision floating-point type has been applied to save the position of a particle because the four-byte single precision is sufficient for small simulations. However, as the number of particles in simulations increases, the position accuracy from single-precision begins to deteriorate. Since the roundoff error of a single precision variable  $\mathcal{A}$  is  $\varepsilon_{\text{roundoff}}(\mathcal{A}) \sim 10^{-7}\mathcal{A}$ , the maximum roundoff error of the single-precision position with respect to the mean particle separation is

$$\varepsilon_{\text{roundoff}} \left( \frac{r_{\max}}{d_{\text{mean}}} \right) \sim 10^{-7} \frac{L_{\text{box}}}{L_{\text{box}}/N_p^{1/3}} \sim 10^{-7} N_p^{1/3}. \quad (1)$$

For example, if the total number of particles is  $6300^3$  as in the HR4, the maximum roundoff position error lies at the level of a few sub-percent of the mean particle separation, or  $\varepsilon_{\text{roundoff}}(r_{\max}/d_{\text{mean}}) \sim 10^{-3}$ .

On the other hand, in the HR4 as well as in the HR2 & 3, we separate the position of a particle ( $\mathbf{r}$ ) into two vectors as

$$\mathbf{r} = \mathbf{L} + \mathbf{d}, \quad (2)$$

where  $\mathbf{L}$  and  $\mathbf{d}$  are the Lagrangian position and displacement from the Lagrangian position of a particle, respectively. We set the particle index by a row-major order in the Lagrangian configuration and, therefore, it does not require additional memory space to compute  $\mathbf{L}$ . Since the displacement of the simulated particle over the entire HR4 simulation run is less than ten times of the mean particle separation ( $d_{\max} \lesssim 10d_{\text{mean}}$ ), the maximum position error in the HR4 is

$$\varepsilon_{\text{roundoff}} \left( \frac{r_{\max}}{d_{\text{mean}}} \right) \sim 10^{-7} \frac{d_{\max}}{d_{\text{mean}}} \sim 10^{-6}. \quad (3)$$

In this way, we significantly enhanced the accuracy of the particle position without using any additional memory space.

### 2.4. Simulation Specifics

The HR4 was performed on the supercomputer of Tachyon-II installed at KISTI (Korea Institute of Science and Technology Information). We used 8,000 CPU cores over 50 straight days from late November in 2013

to early February in 2014. Even with several system glitches over the allocated time period, we succeeded to complete the simulation in about 50 days for the gravitational evolution of  $6300^3$  particles in a periodic cubic box of a side length  $L_{\text{box}} = 3150 h^{-1}\text{Mpc}$ . The starting redshift is  $z_i = 100$ , which is chosen for particles not to overshoot one grid cell spacing (Lukić et al. 2007) in setting the initial conditions. This high initial redshift, combined with 2LPT, ensures an accurate power spectrum and mass function measurement at  $z = 0$  (L'Huillier et al. 2014). The simulation took 2000 steps to reach the final epoch of  $z_0 = 0$ . The mean particle separation is set to  $d_{\text{mean}} = 0.5 h^{-1}\text{Mpc}$  and the corresponding force resolution is  $0.1d_{\text{mean}}$ .

We adopted a standard  $\Lambda\text{CDM}$  cosmology in concordance with WMAP 5-year. This choice of cosmology was made for consistency with various observations including SDSS as well as the previous HRs. Specifically, the matter, baryonic matter, and dark energy densities are  $\Omega_{m,0} = 0.26$ ,  $\Omega_{b,0} = 0.044$ , and  $\Omega_{\Lambda,0} = 0.74$ , respectively. The current Hubble expansion is  $H_0 = 100 h \text{ km/s/Mpc}$ , where  $h = 0.72$ . The amplitude of the initial matter perturbations is scaled for an input bias factor,  $b_8 \equiv 1/\sigma_8 = 1.26$ , where

$$\sigma_8^2 = \frac{1}{2\pi^2} \int k^2 P(k) |W(kR_8)|^2 dk, \quad (4)$$

and  $R_8 \equiv 8h^{-1}\text{Mpc}$ . Here, we used the spherical top-hat filter  $W(x) \equiv 3(x \sin x - \cos x)/x^3$  in  $k$ -space. The particle mass is  $m_p \simeq 9 \times 10^9 h^{-1}\text{M}_\odot$ , and the minimum mass of halos with 30 member particles is about  $M_s \simeq 2.7 \times 10^{11} h^{-1}\text{M}_\odot$ .

Figure 2 shows the evolution of the non-linear matter power spectrum obtained during the simulation run at several redshifts. The dotted lines are the expected linear power spectra, while the solid lines are the simulated matter power spectra at the same redshift. The typical non-linear evolution effect can easily be seen on small scales, where the amplitude of the power spectrum is greater than the linear prediction due to the gravitational clustering.

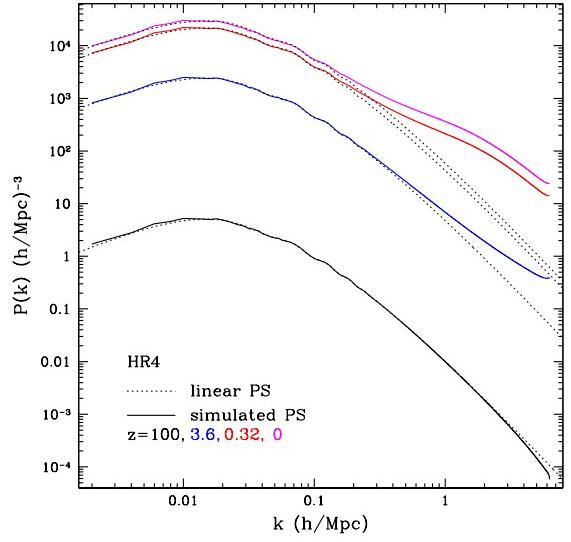
Figure 3 shows a part of the density map of the HR4 at  $z = 0$ , where a cluster develops at the center through the mergers of several neighboring overdensity clumps. One may clearly see void regions (painted in dark blue) with a size of a few tens of  $h^{-1}\text{Mpc}$ . Some overdense blobs are embedded in the connection of multiple filamentary structures.

### 3. OUTPUTS

In this section, we describe the main products of the HR4 simulation. They are available at <http://sdss.kias.re.kr/astro/Horizon-Run4/>.

#### 3.1. Snapshot and Past Lightcone Space Data

We have saved snapshot data of particles at twelve redshifts:  $z = 0, 0.05, 0.1, 0.15, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 1$ , and  $4$ . Each data set contains the particle position, velocity, and eight-byte integer ID index.



**Figure 2.** Matter power spectra from the HR4 simulation (solid) and from linear theory (dotted). Color:  $z = 100$  (black), 3.6 (blue), 0.32 (red), and 0 (magenta).

To generate the past lightcone space data, we put an artificial observer at the origin  $(x, y, z) = (0, 0, 0)$  of the simulation box. At each time step, we calculate the comoving distance from the observer using

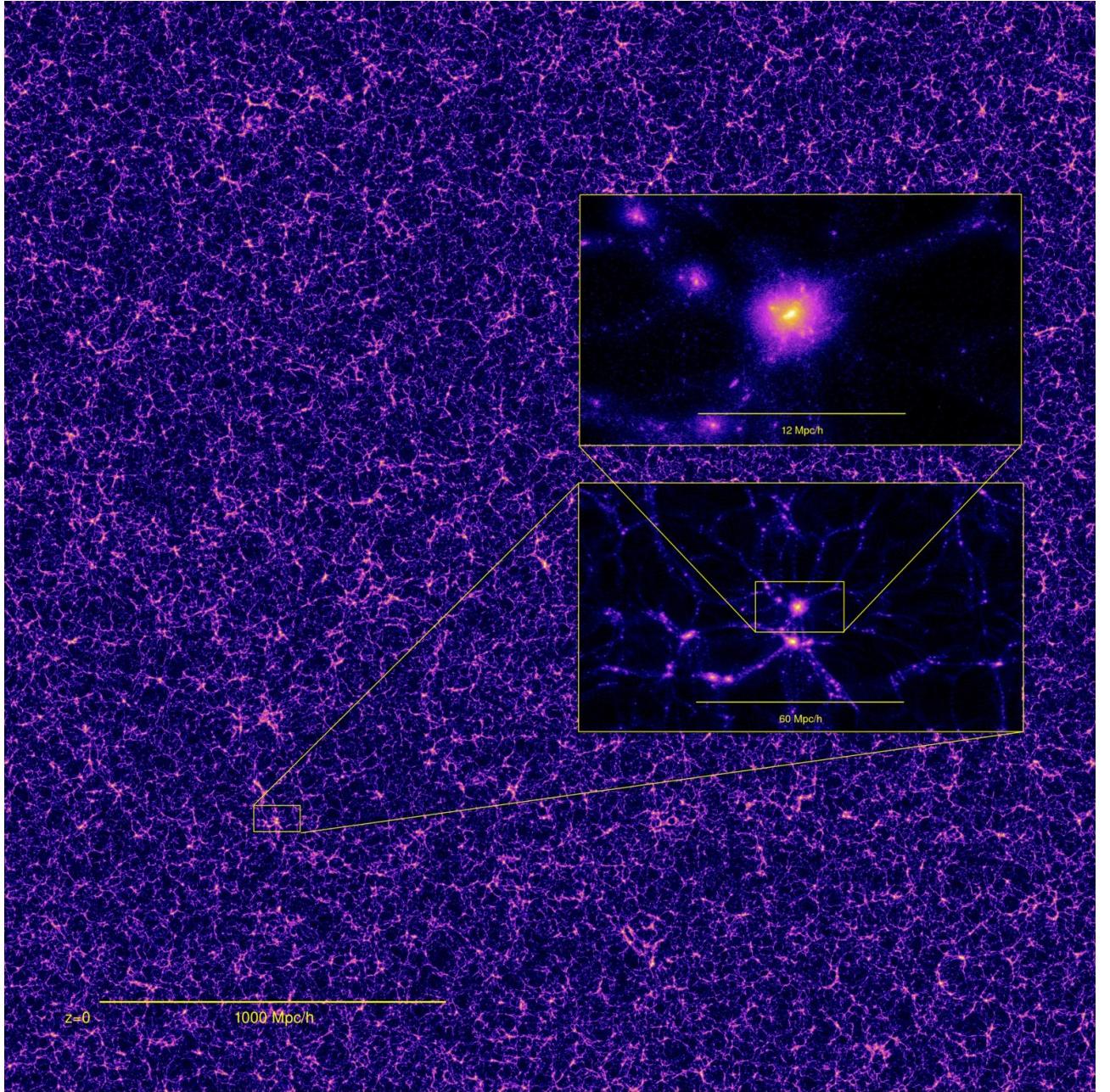
$$d_c = \frac{c}{H_0} \int_0^z \frac{1}{E(z')} dz', \quad (5)$$

where

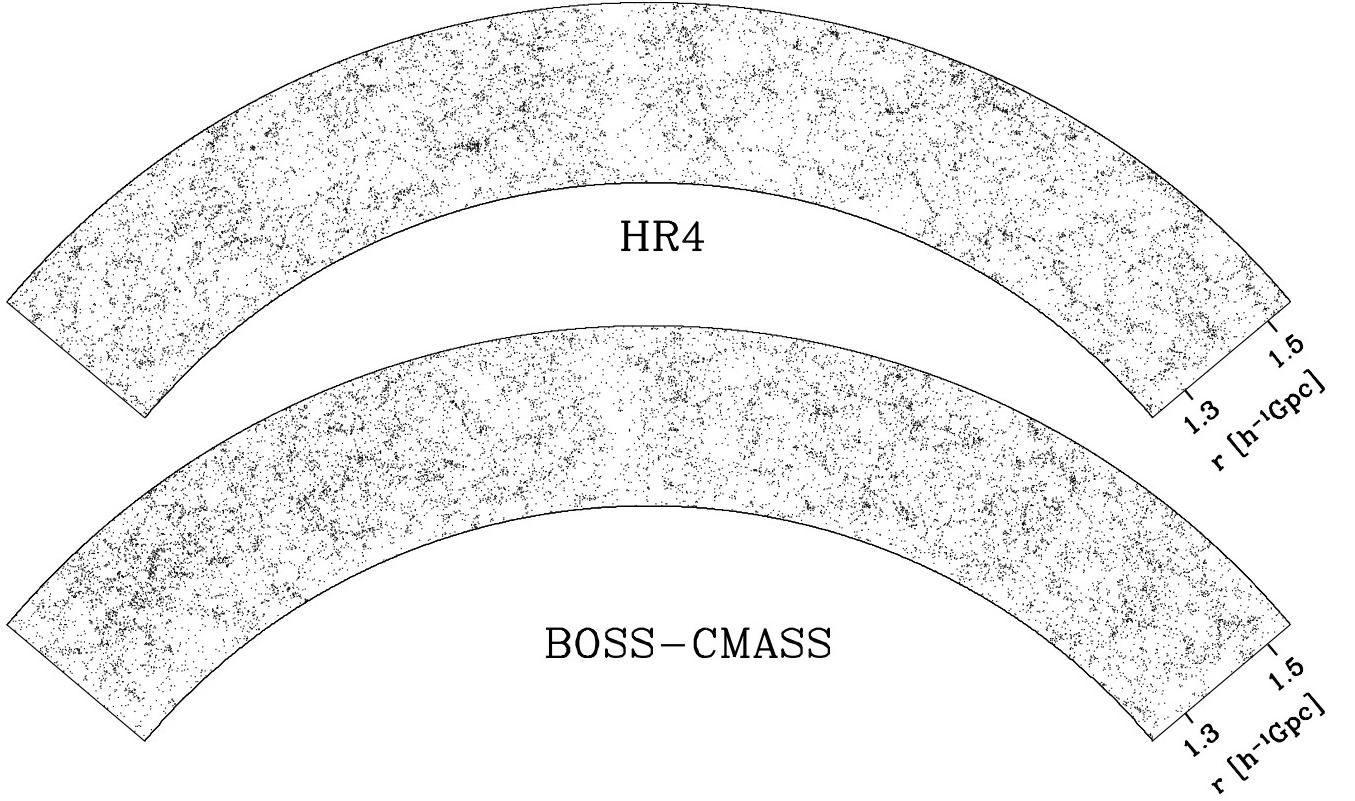
$$E(z) \equiv \sqrt{\Omega_{m,0}(1+z)^3 + (1-\Omega_{m,0}-\Omega_{\Lambda,0})(1+z)^2 + \Omega_{\Lambda,0}}, \quad (6)$$

Then, we search for particles located in a comoving shell, whose inner and outer boundaries at the  $i$ -th step are  $d_{c,i} - \Delta d_{c,i}/2$  and  $d_{c,i} + \Delta d_{c,i+1}/2$ , respectively, where  $\Delta d_{c,i+1} \equiv (d_{c,i+1} - d_{c,i})/2$ . We utilize the periodic boundary conditions by copying the simulation box to extend the all-sky past lightcone space data up to  $r = 3150 h^{-1}\text{Mpc}$ , which corresponds to  $z \simeq 1.5$ .

Due to the finite step size, several undesirable events may be encountered. If a particle crosses the shell boundary between two neighboring time steps, it can be missed or be counted twice in the lightcone space data. Therefore, we set a buffer zone laid upon both sides of the shell. The width of the buffer zone is determined to be equal to the maximum displacement taken by a particle in a time step. Using these buffer zones, we can catch these crossing events. A crossing particle can simultaneously be detected in two contacting shells or two adjacent buffer zones, and we simply merge the duplicated particle by averaging position and velocity in the lightcone space data.



**Figure 3.** Simulation density slice map at  $z = 0$ . High-density regions are painted with bright color. The width of the slice is  $7 h^{-1}\text{Mpc}$ . The two subfigures are arranged for cascaded zoom-in views of a cluster at the center of the box in the bottom part of the figure. We put a scaling bar on the bottom of each panel.



**Figure 4.** *Bottom:* Distribution of galaxies with  $M_r < -22.35$  and  $0.45 < z < 0.6$  in the BOSS-CMASS volume-limited sample (Choi et al. 2015). Galaxies are selected from a strip with  $-2^\circ \leq \text{dec.} \leq 2^\circ$  and  $130^\circ \leq \text{R.A.} \leq 230^\circ$  out of the entire BOSS survey area. *Top:* The ‘galaxies’ in the mock BOSS-CMASS survey performed in the HR4 simulation. The absolute magnitude of the observed galaxies is calculated with the reference redshift of  $z = 0.55$  and is used to produce the sample with a constant number density. The HR4 PSB subhalos are selected as galaxies in accordance with the galaxy-subhalo correspondence model, and the galaxy number density at a given redshift is matched with the observed one by adjusting the low-mass cutoff.

In both the snapshot and past lightcone data, we apply the Ordinary Parallel Friend-of-Friend (OPFOF), a parallel version of FoF code to identify virialized halos. The standard percolation length is simply adopted as  $l_{\text{link}} = 0.2d_{\text{mean}}$ . The halo position and peculiar velocity are given as the average position and velocity of the member particles. Then we apply the physically self-bound (PSB) subhalo finding method (Kim & Park 2006) to identify subhalos embedded in the FoF halo. It employs a negative total energy criterion and spherical tidal boundaries to discard particles from subhalo candidates.

As a representative example, Figure 4 compares the volume-limited galaxy sample from BOSS-CMASS with  $r$ -band magnitude limit  $M_r < -22.35$  at  $0.45 \leq z \leq 0.6$  (*bottom*) with the mock galaxy sample from the HR4 built by the PSB subhalo-galaxy correspondence model (*top*; Kim et al. 2008).

### 3.2. Halo Merger Data

To build the merger trees we detect halos and subhalos at 75 equally-spaced sparse time steps from  $z = 12$  to 0. The step size is set to be comparable to the rota-

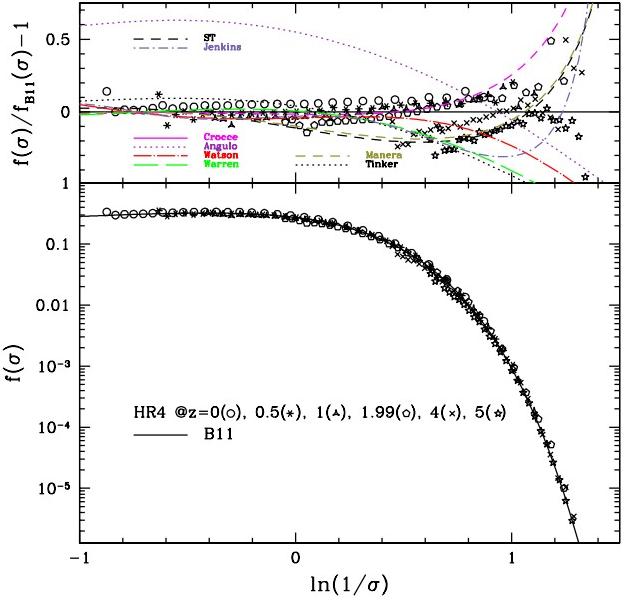
tional period (i.e. dynamical timescale) of Milky-Way-size galaxies. Halo merger trees are then built by tracing the gravitationally most bound member particles (MBPs) of halos. If a halo does not contain any former MBP, we select a new MBP among the member particles of the halo. If one MBP is found, we assume the halo to be a direct descendant of the halo. If the halo hosts multiple former MBPs, we treat the halo as a merger remnant and those ancestor MBPs (or halos) are linked to the remnant creating a halo merger tree. These merger trees will be extensively used to build mock galaxies and to compare with observations (Hong et al. 2015). Of course, due to the halo mass resolution of the HR4 ( $M_s = 2.7 \times 10^{11} h^{-1} M_\odot$ ), we are unable to resolve mergers of sub-Milky-Way-mass (sub)halos.

## 4. PROPERTIES OF FOF HALOS

### 4.1. Multiplicity Function

The multiplicity function is defined as

$$f(\sigma, z) \equiv \frac{M}{\rho_b(z)} \frac{\mathrm{d}n(M, z)}{\mathrm{d} \ln [1/\sigma(M, z)]}, \quad (7)$$



**Figure 5.** Bottom: Multiplicity functions of HR4 FoF halos at  $z = 0, 0.5, 1, 1.99, 4$ , and  $5$ . The solid curve is proposed by Bhattacharya et al. (2011), B11. Top: Fractional deviations of the simulated (symbols) and modeled (lines) multiplicity functions from B11.

where  $n(M, z)$  is the cumulative halo mass function at  $z$ ,  $\rho_b(z)$  is the background matter density, and  $\sigma(M, z)$  is the density fluctuation measured on the mass scale of  $M$ . For a given power spectrum  $P(k)$ , the density fluctuation is estimated as

$$\sigma^2(M, z) = \frac{D_1^2(z)}{2\pi^2} \int k^2 P(k) |W(kR(M, z))|^2 dk, \quad (8)$$

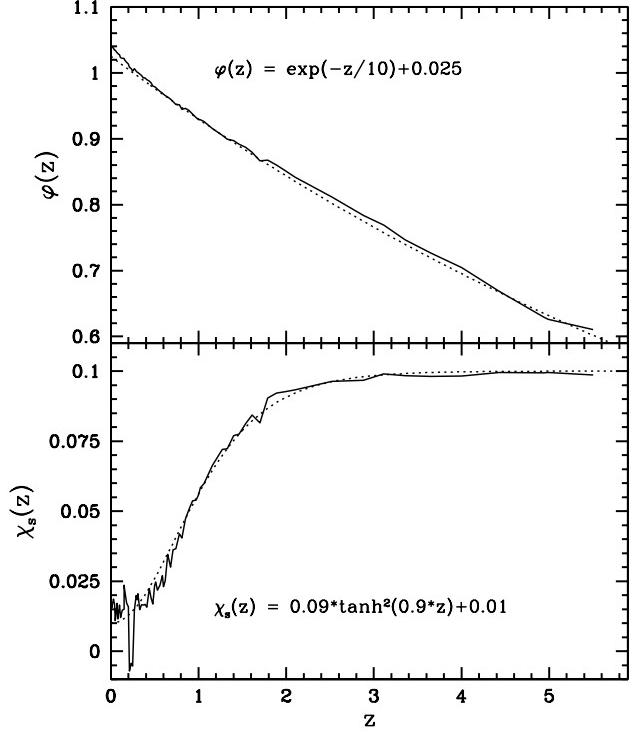
where

$$R(M, z) \equiv \left( \frac{3M}{4\pi\rho_b(z)} \right)^{1/3}, \quad (9)$$

and  $D_1(z)$  is the growing mode of the linear growth factors computed as

$$D_1(z) = \frac{5}{2} \Omega_{m,0} E(z) \int_z^\infty \frac{(1+z')dz'}{E^3(z')} \quad (10)$$

Figure 5 shows the multiplicity function from the HR4 as well as a number of previous fitting models of the multiplicity function (see Table 1 and references therein). Top panel shows the deviations of multiplicity functions with respect to the model described in Bhattacharya et al. (2011), hereafter B11. For clarity, in the cases of Crocce et al. (2010) and Manera et al. (2010) we only show the fitting function obtained at  $z = 0$ . All the previous fitting functions significantly deviate from each other at high mass scales. This may be produced by the exponential cut off producing large noises in fitting the steep slope. From the simulated multiplicity function, we can clearly see the redshift change, and therefore a single functional form may not be sufficient. For large values of  $\ln(1/\sigma)$ , the redshift evolution of



**Figure 6.** Redshift-dependent  $\chi$ - and amplitude-corrections in Equation 11 showing the best fit to the HR4 FoF multiplicity functions. The dotted lines are the analytic fitting functions as shown in Equations (12) and (13).

multiplicity functions is substantial (Lukić et al. 2007), and the overall amplitude seems to increase as the redshift decreases.

We fit the simulated multiplicity function with a variant of the B11 function with an amplitude changing with redshift as

$$f_{\text{Kim}}(\chi_L, z) \equiv \varphi(z) f_{\text{B11}}(\chi_L(M, z) - \chi_s(z)). \quad (11)$$

Here  $\chi_L(M, z) \equiv \sqrt{q}\delta_c/\sigma(M, z)$ , where  $\delta_c = 1.686$  is the density contrast at the collapse epoch in an Einstein-de Sitter universe,  $q$  is a fitting parameter in the B11 function (see Table 1), and  $\chi_s(z)$  and  $\varphi(z)$  are redshift-dependent  $\chi$ - and amplitude-corrections, respectively. The value of  $\delta_c$  in an Einstein-de Sitter Universe is 1.686, and slightly depends on the cosmology. However, for consistency with previous work, we will use this constant value of 1.686 (e.g., Bhattacharya et al. 2011). We fit the simulated multiplicity function with the least-square minimization and obtain the empirical fitting function as

$$\chi_s(z) = 0.09 \tanh^2(0.9z) + 0.01 \quad (12)$$

$$\varphi(z) = \exp\left(-\frac{z}{10}\right) + 0.025. \quad (13)$$

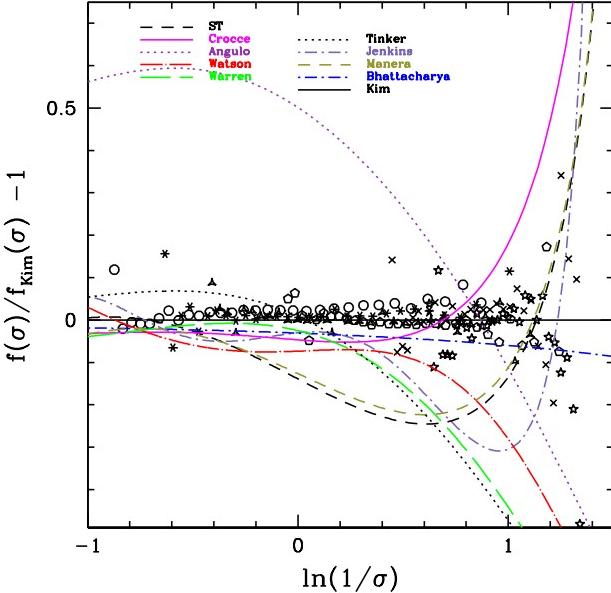
Figure 6 shows the redshift evolution of  $\chi_s(z)$  (left) and  $\varphi(z)$ , respectively. At high redshifts, the HR4 simulation underpopulates halos compared to B11, while the HR4 has more halos than B11 in the recent epoch.

**Table 1**  
Description of fitting models of the multiplicity function

Model	$f(\sigma, z)$	Parameters
Sheth & Tormen (1999)*	$A\sqrt{\frac{2}{\pi}}\chi(1+\chi^{-2p})\exp\left[-\frac{\chi^2}{2}\right]$	$(A, p, q) = (0.3222, 0.3, 0.707, 0.3)$
Jenkins et al. (2001)	$A\exp\left(- \ln\sigma^{-1}+a ^b\right)$	$(A, a, b) = (0.315, 0.61, 3.8)$
Warren et al. (2006)	$A(\sigma^{-a}+b)\exp\left[-\frac{c}{\sigma^2}\right]$	$(A, a, b, c) = (0.7234, 1.625, 0.2538, 1.1982)$
Tinker et al. (2008) <sup>†</sup>	$A(\sigma^{-a}+b)\exp\left[-\frac{c}{\sigma^2}\right]$	$(A, a, b, c) = (0.745, 1.47, 0.250, 1.19)$
Crocce et al. (2010) <sup>†</sup>	$A(\sigma^{-a}+b)\exp\left[-\frac{c}{\sigma^2}\right]$	$(A, a, b, c) = (0.58, 1.37, 0.30, 1.036)$
Manera et al. (2010)* <sup>†</sup>	$A\sqrt{\frac{2}{\pi}}\chi(1+\chi^{-2p})\exp\left[-\frac{\chi^2}{2}\right]$	$(A, p, q) = (0.3222, 0.248, 0.709)$
Bhattacharya et al. (2011)*	$A\sqrt{\frac{2}{\pi}}\chi^r(1+\chi^{-2p})\exp\left[-\frac{\chi^2}{2}\right]$	$(A, p, q, r) = (0.333, 0.807, 0.788, 1.795)$
Angulo et al. (2012)	$A\left(\frac{B}{\sigma}+1\right)^q\exp\left[-\frac{C}{\sigma^2}\right]$	$(A, B, C, q) = (0.201, 2.08, 1.172, 1.7)$
Watson et al. (2013)	$A(\sigma^{-a}+b)\exp\left[-\frac{c}{\sigma^2}\right]$	$(A, a, b, c) = (0.589, 2.163, 0.479, 1.210)$

\*  $\chi \equiv \sqrt{q\delta_c}/\sigma$ , where  $\delta_c = 1.686$  is the density contrast at the collapse epoch in an Einstein-de Sitter universe.

<sup>†</sup> Only the case at  $z = 0$  is given here.

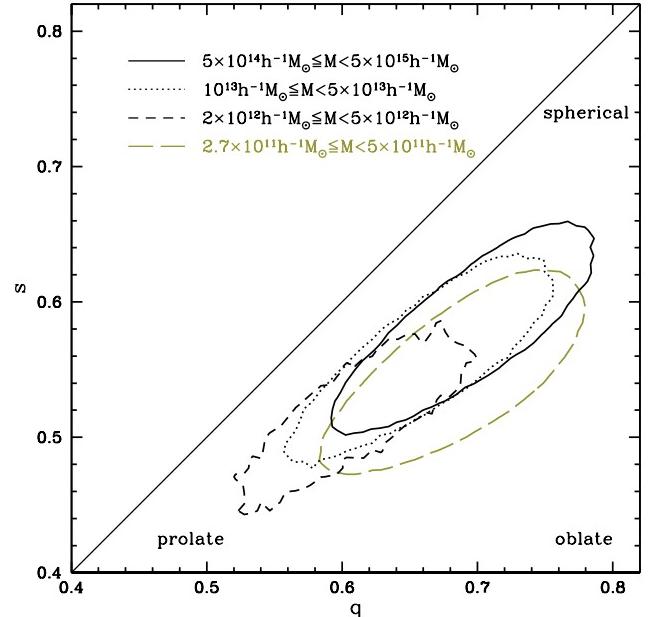


**Figure 7.** Simulated and modeled multiplicity functions with respect to our new fitting model ( $f_{\text{Kim}}(\sigma, z)$ ). Same symbol and color conventions as in Figure 5.

Also, as we move to higher redshift,  $\chi_s(z)$  increases and reaches about 0.098 at very high redshift.

Figure 7 shows the scatter of the simulated multiplicity function (symbols) over our fitting model ( $f_{\text{Kim}}$ ) with various former models (lines) for comparison. While other fitting models match the simulated multiplicity function only at small scales ( $\ln(1/\sigma) < 0.3$ ), our new fitting model agrees with the HR4 on most scales in the redshift range between  $z = 5$  and 0, within a 2.5% fluctuation level.

The origin of the redshift evolution of the multiplic-



**Figure 8.** Shape distributions of FoF halos in various mass samples shown in the  $q$ - $s$  diagram at  $z = 0$ . We mark iso-density contours enclosing 25% halos around peak distributions.

ity function is not clearly known but it might be partly explained by the following arguments. First, even the 2LPT may be somewhat insufficient to generate accurate initial conditions of the simulation. Such effect would be avoided by introducing higher-order approximations (for comparison between the Zel'dovich approximation and 2LPT, see Crocce et al. 2006). However, the target redshifts to measure the halo mass function are sufficiently lower ( $z \lesssim 5$ ; see the discussions made by Tatekawa & Mizuno 2007) compared to the initial epoch of  $z_i = 100$  and, consequently, the redshift evolution may not be caused by numerical transients. Second, it may be due to the limitation of the linear perturbation theory or the linear growth model in calculating  $\sigma(M, z)$  after the nonlinear gravitational clustering begins to enter. The multiplicity function assumes that there is no other redshift dependence than the matter fluctuation,  $\sigma(M, z) = D(z)\sigma(M)$ . But as the non-linear growth becomes significant at lower redshifts, one should consider the effect of non-linear gravitational evolution of density fields.

## 4.2. Halo Shape

### 4.2.1. Structure

The shape tensor  $S_{ij}$  of an FoF halo is defined

$$S_{ij} = \sum_k^{N_m} (x_i^k - \bar{x}_i)(x_j^k - \bar{x}_j), \quad (14)$$

where  $i$  and  $j$  are structural axes,  $\bar{x}$  is the position of the center of the halo mass and  $N_m$  is the number of member particles. The three eigenvalues of the shape tensor  $r_3 \leq r_2 \leq r_1$  are respectively the lengths of the minor, intermediate, and major axes of the corresponding ellipsoid. The prolateness and sphericity of a halo are defined as

$$q \equiv \frac{r_2}{r_1} \quad (15)$$

$$s \equiv \frac{r_3}{r_1}. \quad (16)$$

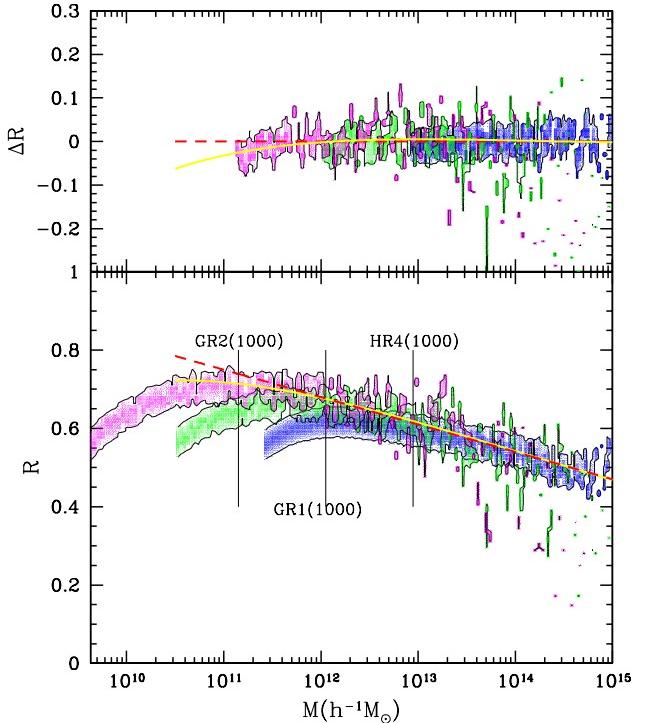
A halo is respectively defined as prolate, oblate, or spherical if

$$q \ll 1, \quad (17)$$

$$q \simeq 1, s \ll 1, \quad (18)$$

$$s \simeq 1. \quad (19)$$

Figure 8 shows the probability distributions of  $(q, s)$  with a contour containing 25% of halos around the peak distribution in four different mass samples of FoF halos. For halo samples more massive than  $2 \times 10^{12} h^{-1} M_\odot$ , we can clearly see that halos become more prolate as the mass increases, in agreement with theoretical predictions (e.g. Rossi et al. 2011). Less massive halos with  $M_s \leq M < 5 \times 10^{11} h^{-1} M_\odot$  have their distribution substantially shifted to the lower-right corner in this diagram, i.e., are more oblate than more massive samples. This may be fully explained by the particle discreteness effect, which will be described in the next section.



**Figure 9.** Resolution dependence of the roundness parameter at  $z = 0$  for HR4 (blue), GR1 (green), and GR2 (magenta) containing 25% of halos around peak probabilities. *Bottom:* Roundness parameter as a function of halo mass. A vertical bar marks the mass scale equivalent to the mass of 1000 particles ( $10^3 m_p$ ) combined. The best-fitting functions of  $R(M)$  from halos with lower-mass cutoff  $10^3 m_p$  ( $R_{1000}^{fit}$ ; yellow) and  $5 \times 10^3 m_p$  ( $R_{5000}^{fit}$ ; red dash) are shown. *Top:* Deviation of the roundness parameter from  $R_{5000}^{fit}$  with respect to the halo mass. Here we do not show the distribution below the mass of  $10^3 m_p$ .

#### 4.2.2. Roundness

We now investigate the FoF halo shape from a different angle. First, we define the roundness as

$$\mathcal{R} \equiv \sqrt{qs} = \sqrt{\frac{r_2 r_3}{r_1^2}}. \quad (20)$$

To measure the resolution effects on halo shape, we ran two additional higher-resolution simulations called Galaxy Run 1 (GR1) and Galaxy Run 2 (GR2). These simulations used  $2048^3$  particles. We employed the same cosmological model but different simulation box sizes ( $L_{\text{box}}^{\text{GR1}} = 512 h^{-1}\text{Mpc}$  &  $L_{\text{box}}^{\text{GR2}} = 256 h^{-1}\text{Mpc}$ ). The mean particle separations of GR1 and GR2 are  $d_{\text{mean}} = 0.25$  and  $0.125 h^{-1}\text{Mpc}$ , respectively, while the corresponding force resolutions are changed to 0.025 and  $0.0125 h^{-1}\text{Mpc}$  accordingly. Therefore, the mass and force resolutions are quite enhanced with respect to the HR4.

Figure 9 shows the distribution of  $\mathcal{R}$  of FoF halos at  $z = 0$ . In each simulation, at scales of  $M \gg 10^3 m_p$ ,  $\mathcal{R}$  tends to be independent of the simulation resolution. On the other hand, at small mass scales ( $M \lesssim 10^3 m_p$ ) each simulation seems to underestimate  $\mathcal{R}$ , probably due to the small number of particles. Hoffmann et al. (2014) examined the discreteness effect on a modeled halo for a given shape and found that the required number of particles should not be less than 1000 for a reliable shape determination (see Fig. A2 in their paper).

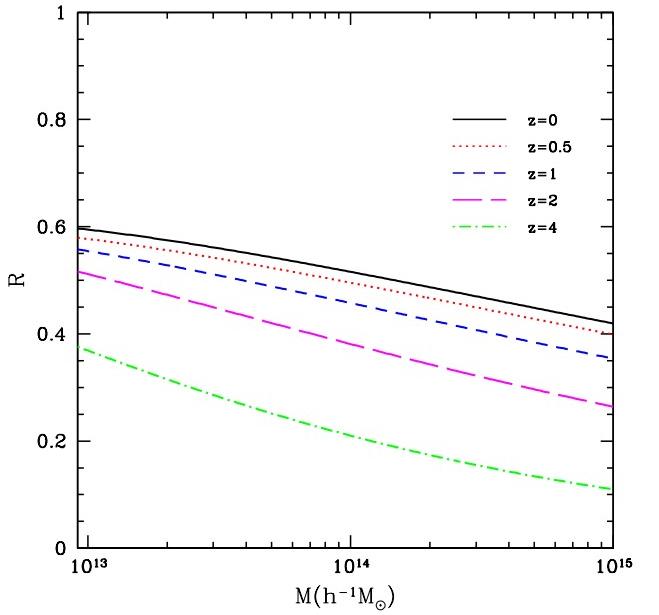
From our three simulations we found a fitting formula of the roundness parameter as a function of the halo mass for massive halos with  $M \geq 10^3 m_p$ :

$$\begin{aligned} \mathcal{R}_{1000}^{\text{fit}}(M) &= a_{\mathcal{R}} \log_{10} \left( \frac{M}{10^7 h^{-1}\text{M}_\odot} \right) \\ &\times \exp \left[ -b_{\mathcal{R}} \log_{10} \left( \frac{M}{10^7 h^{-1}\text{M}_\odot} \right) \right], \end{aligned} \quad (21)$$

where  $(a_{\mathcal{R}}, b_{\mathcal{R}}) = (0.55, 0.28)$  (yellow line in Figure 9). One should note that this fitting is only valid for  $M \gtrsim 1.4 \times 10^{11} h^{-1}\text{M}_\odot$ , which is set by the combined mass of 1000 particles of the GR2. We do not find any turn-around mass scale of  $\mathcal{R}$  in the available mass range in this study. If we only consider halos with  $M \geq 5 \times 10^3 m_p$ , the distribution of  $\mathcal{R}$  follows

$$\mathcal{R}_{5000}^{\text{fit}}(M) = A_{\mathcal{R}} \log_{10} \left( \frac{M}{10^{12} h^{-1}\text{M}_\odot} \right) + B_{\mathcal{R}}, \quad (22)$$

where  $(A_{\mathcal{R}}, B_{\mathcal{R}}) = (-0.07, 0.68)$  (red dashed line in Figure 9). Similar to the case of  $\mathcal{R}_{1000}^{\text{fit}}$ , the above fitting is only valid for  $M \gtrsim 7 \times 10^{11} h^{-1}\text{M}_\odot$ . Both  $\mathcal{R}_{1000}^{\text{fit}}$  and  $\mathcal{R}_{5000}^{\text{fit}}$  describe well the change of  $\mathcal{R}$ , but they diverge below  $M = 7 \times 10^{11} h^{-1}\text{M}_\odot$ , the mass scale of about  $5 \times 10^3 m_p$  of GR2. Therefore, we may need a simulation with a higher mass resolution than GR2 to see which fitting formula describes the roundness parameter of low-mass halos.



**Figure 10.** Change of peak position in the  $\mathcal{R}$  distribution at several redshifts in the HR4. The lower bound of the  $x$ -axis corresponds to  $10^3 m_p$ .

Figure 10 shows the redshift evolution of  $\mathcal{R}$  in the HR4. At higher redshift, halos tend to have a smaller value of  $\mathcal{R}$  for a given mass. However, it is important to note that this tendency does not guarantee a possible shape evolution of virialized halos because halos also grow in mass with time.

#### 4.3. Halo Orientations

In this section, we study the angle between the halo rotational and structural axes. The directional angle between them is calculated as

$$\theta_i = \cos^{-1} |\hat{\mathbf{J}} \cdot \hat{\mathbf{r}}_i|, \quad (23)$$

where  $\hat{\mathbf{J}}$  is the normalized rotational axis and  $\hat{\mathbf{r}}_i$  is the unit vector in the direction of the structural axis  $i$ . We define the probability distribution function of the directional angles

$$p'(\theta) \equiv \frac{dP(\theta)}{d \cos \theta}, \quad (24)$$

where  $P(\theta)$  is the cumulative probability of a directional angle greater than  $\theta$ . Then, for a random orientation  $p'(\theta)$  is uniform over the angle of  $0^\circ \leq \theta < 90^\circ$ .

Figure 11 shows the relations between the rotation and halo axes as a function of halo mass. The rotational axis tends to be orthogonal to the major axis (bottom panel), which means that halos tend to swing around their major axis. Moreover, from the upper two panels, it can be noted that the rotational axis is more aligned with the minor axis than the intermediate axis. This alignment becomes stronger as the halo mass increases. In addition, we find that this tendency still holds for low-mass halos with  $M \lesssim 10^3 m_p$ , implying that the halo rotation is less affected by the mass resolution limit than the halo shape.

## 5. TWO-POINT CORRELATION FUNCTION

In this section, we implement the effects of redshift-space distortions on the clustering of mock galaxies and measure the change of clustering in the radial ( $\pi$ ) and tangential ( $\sigma$ ) directions. In the 3-dimensional space, the radial separation between two points ( $\mathbf{r}_1$  and  $\mathbf{r}_2$ ) is defined as

$$\pi \equiv \frac{|\mathbf{d}_{12} \cdot \mathbf{R}_{12}|}{|\mathbf{R}_{12}|}, \quad (25)$$

where  $\mathbf{R}_{12} \equiv (\mathbf{r}_1 + \mathbf{r}_2)/2$  and  $\mathbf{d}_{12} \equiv \mathbf{r}_1 - \mathbf{r}_2$ . The tangential distance between them is simply obtained with

$$\sigma = \sqrt{\mathbf{d}_{12} \cdot \mathbf{d}_{12} - \pi^2}. \quad (26)$$

The correlation function of a point set can easily be calculated by Hamilton's method (Hamilton 1993):

$$\xi(\sigma, \pi) = \frac{\mathbf{DD}(\sigma, \pi) \mathbf{RR}(\sigma, \pi)}{\mathbf{DR}(\sigma, \pi)^2} - 1, \quad (27)$$

Here,  $\mathbf{DD}$  is the number of pairs of real points,  $\mathbf{DR}$  is the number of cross pairs between the real and random points, and  $\mathbf{RR}$  is the number of pairs of random points at the two-dimensional separations of  $\sigma$  and  $\pi$ .

We use the PSB subhalo catalog from the HR4 snapshot at  $z = 0$  as our mock galaxy sample. By adopting the far-field approximation and using the periodic boundary condition of the HR4 simulation, we produce the redshift-space distortion in the  $x$ -direction,

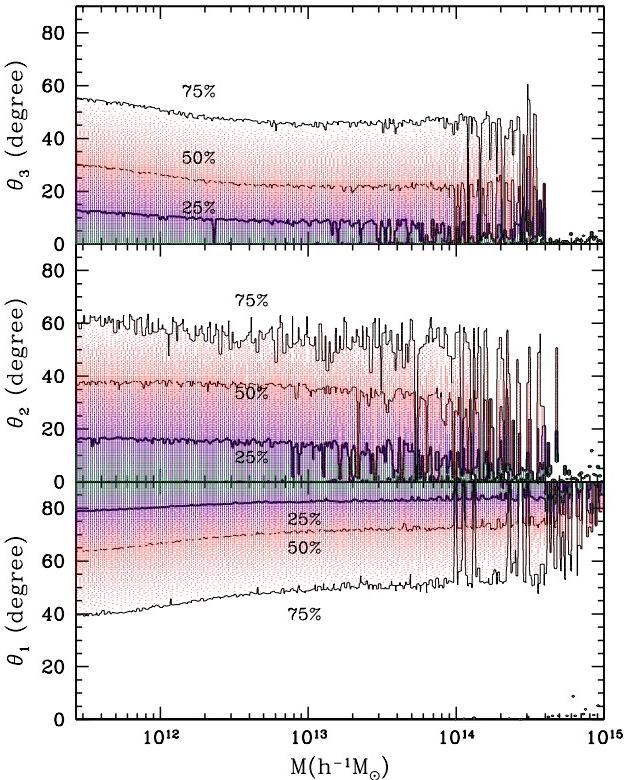
$$x' = x + \frac{v_x}{H_0}, \quad (28)$$

where  $H_0$  is the Hubble parameter at  $z = 0$  and  $v_x$  is the peculiar velocity along the  $x$ -axis. Since we adopt the far-field approximation,  $\pi$  is the position difference in the  $x$ -axis and  $\sigma$  is the separation in the  $y$ - $z$  plane. We then construct a mass-limited mock galaxy sample with PSB subhalos satisfying  $M \geq 2.60 \times 10^{12} h^{-1} M_\odot$ . The average number density of the mass-limited PSB subhalo sample is  $\bar{n} = 1.48 \times 10^{-3} h^3 \text{Mpc}^{-3}$ , which is comparable to the number density of the volume-limited sample of the SDSS Main galaxies with absolute magnitude limit of  $M_r - 5 \log_{10} h < -21$  (Choi et al. 2010).

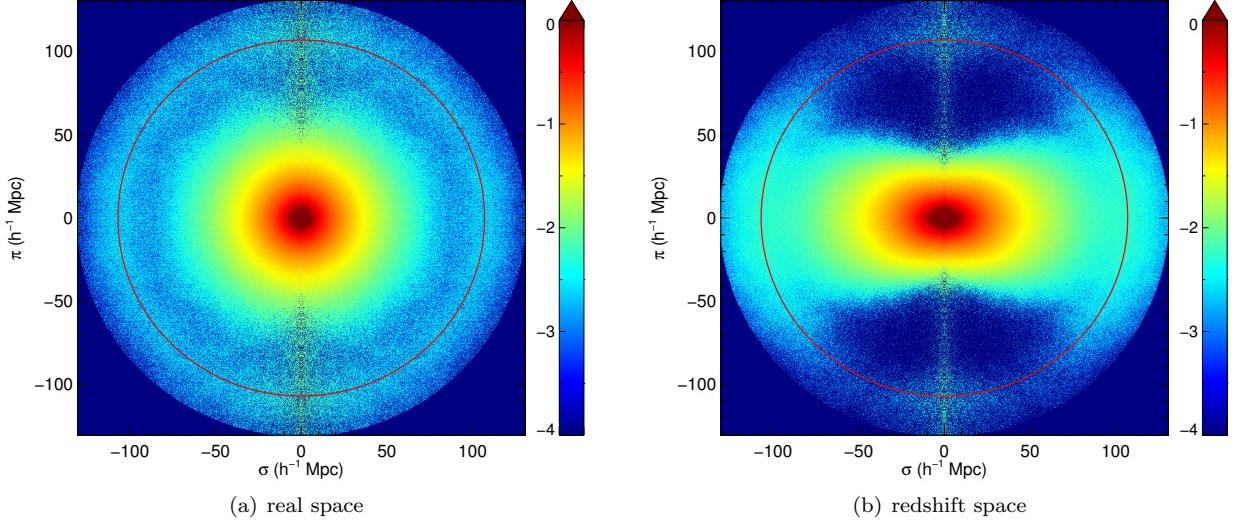
Figure 12 shows the effects of redshift-space distortions on the correlation map. The left panel shows the correlation of our mock galaxy sample in real space while the effects of redshift-space distortion are applied in the right panel. The shape of  $\xi(\pi, \sigma)$  is distorted along the line-of-sight (LoS, or in the  $\pi$ -direction). At the very center, the finger-of-god effect can be seen as spikes stretching along the  $\pi$ -direction (for a better view around the center, see Figure 13). On the other hand, on larger scales, the correlation function along the LoS contracts to the smaller scale.

The position of the BAO peak in real space can be estimated from the linear correlation function

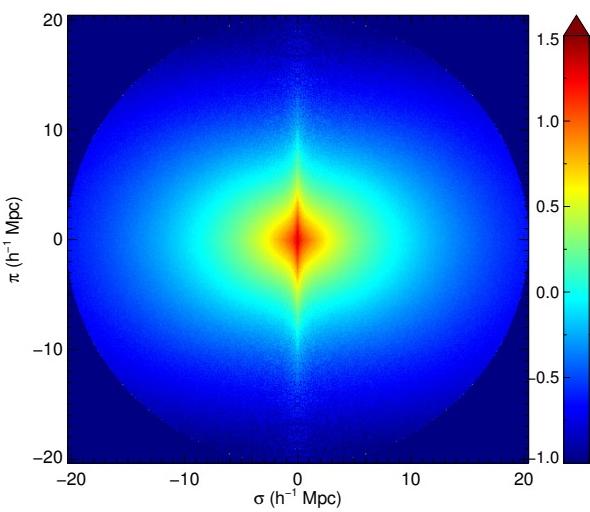
$$\xi_{\text{linear}}(r) \equiv \frac{1}{2\pi^2} \int k^2 P(k) \frac{\sin(kr)}{kr} dk. \quad (29)$$



**Figure 11.** Orientations of the rotational axis with respect to the major (bottom), intermediate (middle), and minor (top panel) axes at  $z = 0$ . Probability contours are drawn around the peak position at each mass bin enclosing 25%, 50%, and 75% of halos respectively. Most halos are positioned at  $\theta_1 \simeq 90^\circ$ ,  $\theta_2 \simeq 0^\circ$ , and  $\theta_3 \simeq 0^\circ$ .



**Figure 12.** Correlation functions of mock galaxies measured without (*left*) and with (*right*) redshift-space distortion effects. The radius of each circular region is  $130 h^{-1}$  Mpc, and the solid circle marks the BAO peak position ( $r_{\text{peak}} \simeq 107 h^{-1}$  Mpc). The color bar marks the correlation in logarithmic spacing.



**Figure 13.** Same as the right panel of Figure 12, but zoomed in to clearly show the finger-of-god effect.

For the WMAP 5-year standard  $\Lambda$ CDM cosmology, the BAO peak in real space is located at  $r_{\text{peak}} \simeq 107 h^{-1}$  Mpc, shown as a solid circle in Figure 12.

Figure 14 shows the two-point correlation functions for different values of the directional cosine to the LoS direction  $\mu$  in real space (*top*) and redshift space (*bottom panel*). In real space, the correlation function around the BAO peak is independent of the directional angle ( $\theta$ ) because of the isotropic distribution. On the other hand, the correlation functions measured in redshift space are increased as  $\theta$  increases, because galaxy pairs are stretched along the LoS. It is worth to note that the BAO peak in the tangential direction ( $\theta = 90^\circ$ ) cannot be detected. Moreover, the correlation functions along the LoS has a peak with a height nearly zero while correlation functions for  $\theta < 30^\circ$  are less than zero on scales below the peak position.

Figure 15 shows the average correlation function over the directional cosine,

$$\xi(r) \equiv \int_0^1 \xi(r, \mu) d\mu. \quad (30)$$

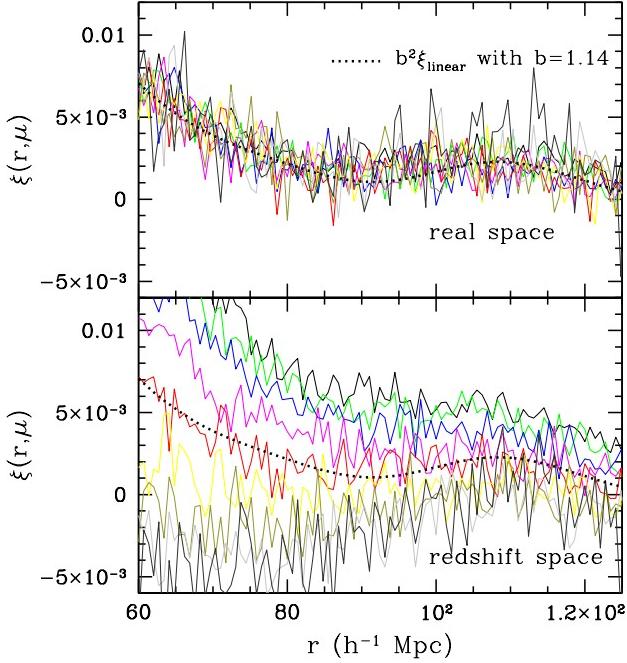
In both real and redshift spaces, the BAO peak from the HR4 is broadened and shifted toward small scales compared to a simple estimation of the linear correlation function of biased objects

$$\xi_{\text{linear}, \text{bias}}(r) = b^2 \xi_{\text{linear}}(r), \quad (31)$$

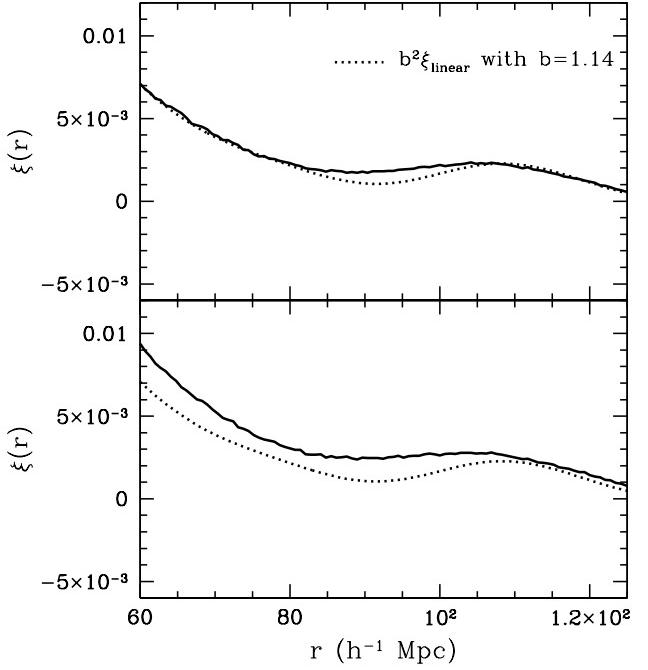
where  $b = 1.14$  is the bias factor. This is due to the nonlinear gravitational evolution of galaxies. In redshift space, the BAO peak is further broadened, and it is hard to clearly find the position of the BAO peak.

## 6. MASS ACCRETION HISTORY

We use merger trees to study the mass accretion history of halos in several mass samples. We define the mass



**Figure 14.** Correlation functions between PSB halo pairs separated along the directions of  $\theta = 10, 20, 30, 40, 50, 60, 70, 80$ , and  $90$  degrees in real space (*top*) and redshift-distorted space (*bottom panel*). In the bottom panel, the top-most line is the correlation of  $\theta = 80^\circ$ , and the correlation function increases as  $\theta$  decreases. The dotted line is the linear prediction with a bias factor  $b = 1.14$ . As the direction cosine ( $\mu \equiv \cos \theta$ ) increases, the noise of the correlation functions is decreasing because the number of pairs along the given direction increases ( $N_{\text{pair}} \propto \mu$ ). The correlation functions along  $\theta = 90^\circ$  are not shown due to big noise.



**Figure 15.** Correlation function averaged over the directional cosine. The thick solid line is the averaged value of the correlation functions in real space (*top*) and redshift space (*bottom panel*). The dotted line is the linear prediction with bias  $b = 1.14$ .

accumulation history as

$$\Psi(M_0, z) \equiv \frac{M(z)}{M_0}, \quad (32)$$

where  $M_0$  is the final halo mass at  $z = 0$ . The half-mass epoch ( $z_{1/2}$ ) is defined as the time when  $M(z_{1/2}) = M_0/2$ . We measure the evolution of the halo mass along the major descendant trees and show the results in Figure 16. It can be seen that the half-mass redshift tends to decrease as the final halo mass increases. For example, low-mass halos with  $10^{12} \leq M_0/h^{-1}\text{M}_\odot < 3 \times 10^{12}$  tend to have their half-mass around  $z_{1/2} \simeq 1$  on average, while more massive halos with  $10^{14} \leq M_0/h^{-1}\text{M}_\odot < 5 \times 10^{14}$  tend to have a later half-mass epoch  $z_{1/2} \simeq 0.5$ . This result is consistent with the observations that galaxy clusters formed relatively recently (in terms of the epoch when a cluster obtains half of the current mass) while individual satellite galaxies seem to form at relatively higher redshift.

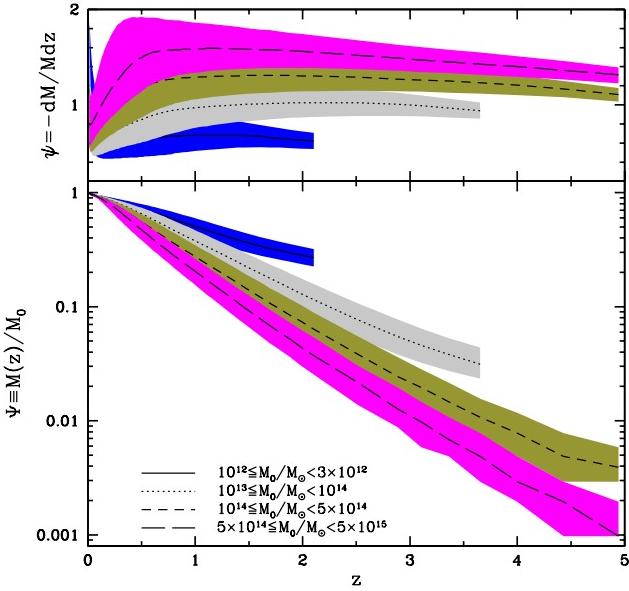
We empirically fit the log-linear function of redshift to  $\Psi(z)$  as

$$\Psi(M_0, z) = \exp [-\psi(M_0)z], \quad (33)$$

and found a best-fit relation as

$$\psi(M_0) \simeq 0.32 \log_{10} \left( \frac{M_0}{10^{12} h^{-1}\text{M}_\odot} \right) + 0.56. \quad (34)$$

This fitting function reproduces the distribution for  $10^{12} \leq M/h^{-1}\text{M}_\odot < 5 \times 10^{15}$  quite well at an early



**Figure 16.** Evolution of halo mass with redshift for several mass samples. In the top panel, we show the change of  $\psi$  with redshift, while  $\Psi$  is shown in the bottom panel. Lines and shaded regions mark the mean and  $1\sigma$  distributions of mass history. For each sample, we cut the data below the mass resolution of the simulation.

epoch ( $z \gtrsim 0.7$ ). On the other hand, at low redshifts ( $z \lesssim 0.7$ ) the accelerated expansion driven by dark energy begins to overpower the gravitational attraction and, therefore, halo mergers are suppressed. The sharp increase of  $\psi$  near the current epoch is caused by a numerical noise. Note that Dekel et al. (2013) also found an exponential form of mass growth, although their mass accretion rate ( $\psi(10^{12} h^{-1} M_\odot) = 0.76$ ) is slightly higher than ours ( $\psi(10^{12} h^{-1} M_\odot) = 0.56$ ).

We want to point out that the specific mass accretion rate per unit redshift interval, defined as

$$\left| \frac{dM}{Mdz} \right| = \left| \frac{M_0}{M} \frac{d\Psi}{dz} \right| \approx \psi(M_0), \quad (35)$$

is roughly constant with redshift and depends only on the current sample mass. Then the specific mass accretion rate per unit physical time can be calculated as

$$\Upsilon_M(M_0, z) \equiv \left| \frac{dM}{Mdt} \right| \quad (36)$$

$$= \frac{\psi(M_0)}{H_0} \frac{E(z)}{1+z}. \quad (37)$$

We now introduce a star formation efficiency, which is defined as the ratio of mass accretion rates between the stellar and total masses of halos as

$$b_\star(M_0, z) \equiv \frac{\Upsilon_\star}{\Upsilon_M}, \quad (38)$$

where  $\Upsilon_\star \equiv dM_\star/M_\star dt$  is the specific stellar mass accretion rate. As a simple case, we assume that the

stellar mass evolution of a halo is fully determined by the evolution of its total mass. In this case, the spectral indices of stellar mass-to-total mass, stellar mass accretion rate-to-stellar mass, and star formation efficiency-to-total mass, respectively defined as

$$\gamma(M_0, z) \equiv \frac{d \ln M_\star}{d \ln M} \quad (39)$$

$$\beta(M_0, z) \equiv \frac{d \ln \Upsilon_\star}{d \ln M_\star} \quad (40)$$

$$\epsilon(M_0, z) \equiv \frac{d \ln b_\star}{d \ln M}, \quad (41)$$

are fully determined by the redshift and the final halo mass. By applying the galaxy-subhalo correspondence model to relate between halo mass and galaxy luminosity, Kim et al. (2008) showed that stellar luminosity (or stellar mass if a constant  $M_\star/L_\star$  is assumed) shows a good relation to the halo mass with a power-law index  $\gamma \sim 0.5$  for the SDSS main galaxy sample when  $M \gtrsim 5 \times 10^{11} h^{-1} M_\odot$ . A similar slope was reported by Kravtsov et al. (2014) from the BCG samples. On the other hand, Abramson et al. (2014) reported  $\beta \sim -0.3$  for SDSS DR7 galaxies with  $9.5 \leq \log_{10}(M_\star/h^{-1} M_\odot) \leq 11.5$ .

The spectral index of the stellar mass accretion rate-to-total mass can be expressed as a combination of the above spectral indices:

$$\eta(M_0, z) \equiv \frac{d \ln \Upsilon_\star}{d \ln M} \quad (42)$$

$$= \beta(M_0, z) \gamma(M_0, z) \quad (43)$$

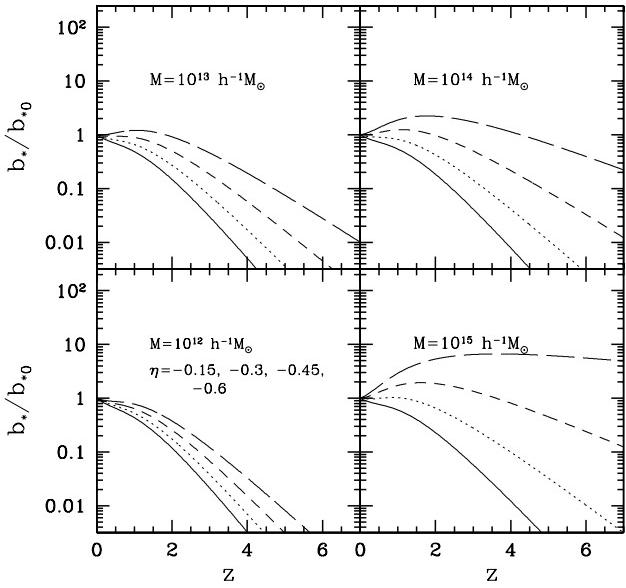
$$= \epsilon(M_0, z) + \frac{1}{\psi(M_0)} \left[ \frac{d \ln E(z)}{dz} - 1 \right], \quad (44)$$

where  $E(z)$  was defined in Equation (6). The effect of the parameters on the relative star formation efficiency is shown in Figure 17. As can be seen from the figure, the relative star formation efficiency is higher, or the  $\eta$  is getting smaller for more massive halos.

## 7. SUMMARY

We ran a new cosmological  $N$ -body simulation called the Horizon Run 4 (HR4) simulation. By adopting a standard  $\Lambda$ CDM cosmology in concordance with WMAP 5-year observations, the HR4 simulates a periodic cubic box of a side length,  $L_{\text{box}} = 3150 h^{-1} \text{Mpc}$  with  $6300^3$  particles. With its wide range of mass and length scales, the HR4 can provide the cosmology community with a competitive data set for the study of cosmological models and galaxy formation in the context of large-scale environments.

The main products of the HR4 are as follows. First, we saved the snapshot data of the particles within the whole simulation box at 12 different redshifts from  $z = 4$  to 0. We also built a past lightcone space data of particles that covers the all-sky up to  $z \simeq 1.5$ . They can be used to study the evolution of the gravitational potential and the genus topology as well as large-scale weak lensing analysis. Moreover, we constructed the



**Figure 17.** Relative star formation efficiency scaled with the current efficiency,  $b_{*0}$ . Clockwise from the lower-left panel, the halo masses are  $M_0 = 10^{12} h^{-1} M_\odot$ ,  $M_0 = 10^{12} h^{-1} M_\odot$ ,  $M_0 = 10^{13} h^{-1} M_\odot$ , and  $M_0 = 10^{15} h^{-1} M_\odot$ , respectively. In the legend, we list the values of  $\eta$  from the bottom curve.

merger trees of Friend-of-Friend halos from  $z = 12$  to 0 with their gravitationally most bound member particles. They can be used to study galaxy formation and bridge the gap between theoretical models and observed galaxy distributions.

We tested the HR4 in various aspects, including the mass/shape/spin distributions of FoF halos, two-point correlation functions of physically self-bound subhalos, and mass evolution of FoF halos. The results of our test are summarized as follows:

1. We found that the abundance of massive FoF halos in the HR4 is substantially different from various fitting functions given in the previous literature. We also found strong evidence for a redshift dependence of the mass function. We proposed a new fitting formula of the multiplicity function that reproduces the redshift changes of amplitude and shape of multiplicity functions within about 5 % errors.
2. We confirmed the finding of previous studies that FoF halos tend to rotate around the minor axis.
3. The two-point correlation function measured in real space is isotropic. However, due to the non-linear evolution of galaxies, the location of the baryonic acoustic oscillation peak is shifted toward smaller scale than the prediction from the linear correlation function. On the other hand, in redshift space the BAO peak can be seen only in the two-point correlation function along the perpendicular direction, with a much-broadened width and increased height. We emphasized that it is im-

portant to use massive simulation data to study the non-linear evolution of BAO features and the connection between observations and cosmological models.

4. We found that more massive halos tend to have steeper mass histories, and the mass accretion rate per unit redshift is roughly constant during early epoch before dark energy domination. By adopting simple power-law models for the stellar mass and star formation efficiency, we found that massive halos tend to have a higher star formation efficiency.

All aforementioned main products of the HR4 are available at <http://sdss.kias.re.kr/astro/Horizon-Run4/>.

## ACKNOWLEDGMENTS

This work was supported by the Supercomputing Center/Korea Institute of Science and Technology Information with supercomputing resources including technical support (KSC-2013-G2-003). The authors thank Korea Institute for Advanced Study for providing computing resources (KIAS Center for Advanced Computation) for this work. The authors also thank the referee, Graziano Rossi, for the thorough review and constructive suggestions that lead to an improvement of the paper.

## REFERENCES

- Abramson, L. E., Kelson, D. D., Dressler, A., Poggianti, B., Gladders, M. D., Oemler, A. Jr., & Vulcani, B. 2014, The Mass-independence of specific star formation rates in galactic disks, *ApJ*, 785L, 36
- Ahn, J., Kim, J., Shin, J., Kim, S. S., & Choi, Y.-Y. 2014, Halo spin parameter in cosmological simulations, *Journal of Korean Astronomical Society*, 47, 77
- Angulo, R. E., Springel, V., White, S. D. M., Jenkins, A., Baugh, C. M. & Frenk, C. S. 2012, Scaling relations for galaxy clusters in the Millennium-XXL simulation, *MNRAS*, 426, 2046
- Athanassoula, E., Vozikis, C. L., & Lambert, J. C. 2001, Relaxation times calculated from angular deflections, *A&A*, 376, 1135
- Baugh, C. M. 2006, A primer on hierarchical galaxy formation: the semi-analytical approach, *Rep. Prog. Phys.*, 69, 310
- Berlind, A. A. & Weinberg, D. H. 2002, The halo occupation distribution: toward an empirical determination of the relation between galaxies and mass, *ApJ*, 575, 587
- Bhattacharya, S., Heitmann, K., White, M., Lukic, Z., Wagner, C., & Habib, H. 2011, Mass function predictions beyond CDM, *ApJ*, 732, 122
- Blanton, M. R. & Berlind, A. A. 2007, What aspects of galaxy environment matter?, *ApJ*, 664, 791
- Bouchet, F. R., Colombi, S., Hivon, E., & Juszkiewicz, R. 1995, Perturbative Lagrangian approach to gravitational instability, *A&A*, 296, 575
- Boylan-Kolchin, M., Ma, C.-P., & Quataert, E. 2008, Dynamical friction and galaxy merging time-scales, *MNRAS*, 383, 93
- Choi, Y.-Y., Han, D.-H., Kim, S. S. 2010, Korea Institute for Advanced Study Value-Added Galaxy Catalog, *JKAS*, 43, 191
- Choi, Y.-Y. et al. 2015, in preparation

- Choi, Y.-Y., Kim, J., Rossi, G., Kim, S. S., & Lee, J.-E. 2013, Topology of luminous red galaxies from the Sloan digital Sky Survey, *ApJS*, 209, 19
- Cole, S., Aragon-Salamanca, A., Frenk, C. S., Navarro, J. F., & Zepf, S. E. 1994, A recipe for galaxy formation, *MNRAS*, 271, 781
- Crocce, M., Pueblas, S., & Scoccimarro, R. 2006, Transients from initial conditions in cosmological simulations, *MNRAS*, 373, 369
- Crocce, M., Fosalba, P., Castander, F. J., & Gaztanaga, E. 2010, Simulating the Universe with MICE: the abundance of massive clusters, *MNRAS*, 403, 1353
- De Lucia, G., Kauffmann, G., & White, S. D. M. 2004, Substructures in cold dark matter haloes, *MNRAS*, 349, 1101
- Dekel, A., Zolotov, A., Tweed, D., et al. 2013, Toy models for galaxy formation versus simulations, *MNRAS*, 435, 999
- Dubinski, J. 1996, A parallel tree code, *New Astronomy*, 1, 133
- Dubinski, J., Kim, J., Park, Changbom, & Humble, R. 2004, GOTPM: a parallel hybrid particle-mesh treecode, *New Astronomy*, 9, 111
- Dunkley, J., Komatsu, E., Nolta, M. R., et al. 2009, Five-Year Wilkinson Microwave Anisotropy Probe observations: likelihoods and parameters from the WMAP data, *ApJSS*, 180, 306
- Eisenstein, D. J. & Hu, W. 1998, Baryonic features in the matter transfer function, *ApJ*, 496, 605
- Gao, L., De Lucia, G., White, S. D. M., & Jenkins, A. 2004, Galaxies and subhaloes in CDM galaxy clusters, *MNRAS*, 352, L1
- Guo, Q., White, S. D. M., Boylan-Kolchin, M., De Lucia, G., Kauffmann, G., Lemson, G., Li, C., Springel, V., & Weinmann, S. 2011, From dwarf spheroidals to cD galaxies: simulating the galaxy population in a CDM cosmology, *MNRAS*, 413, 101
- Guo, Q. & White, S. 2014, Numerical resolution limits on subhalo abundance matching, *MNRAS*, 437, 3228
- Hamilton, A. J. S. 1993, Toward better ways to measure the galaxy correlation function, *ApJ*, 417, 19
- Hoffmann, K., et al. 2014, Subhaloes gone Notts: subhaloes as tracers of the dark matter halo shape, *MNRAS*, 442, 1197
- Hong, S. E., Park, C., & Kim, J. 2015, The most bound halo particle - galaxy correspondence model of galaxy distribution: implementation of merger timescale, submitted to *ApJ*
- Huang, S., Dubinski, J., & Carlberg, R. G. 1993, Orbital deflections in N-body systems, *ApJ*, 404, 73
- Hwang, J.-C., Noh, H., & Gong, J.-O. 2012, Second-order solutions of cosmological perturbation in the matter-dominated era, *ApJ*, 752, 50
- Jenkins, A., Frenk, C. S., White, S. D. M., Colberg, J. M., Cole, S., Evrard, A. E., Couchman, H. M. P. & Yoshida, N. 2001, The mass function of dark matter haloes, *MNRAS* 321, 372
- Jenkins, A. 2010, Second-order Lagrangian perturbation theory initial conditions for resimulations, *MNRAS*, 403, 1859
- Jiang, C. Y., Jing, Y. P., Faltenbacher, A., Lin, W. P., & Li, C. 2008, A fitting formula for the merger timescale of galaxies in Hierarchical clustering, *ApJ*, 675, 1095
- Kauffmann, G., Nusser, A., & Steinmetz, M. 1997, Galaxy formtaion and large-scale bias, *MNRAS*, 286, 795
- Kauffmann, G., White, S. D. M., Heckman, T. M., Ménard, B., Brinchmann, J., Charlot, S., Tremonti, G., Brinkmann, J. 2004, The environmental dependence of the relations between stellar mass, structure, star formation and nuclear acitivity in galaxies, *MNRAS*, 353, 713
- Kim, Y.-R., Choi, Y.-Y., Kim, S. S., Kim, K.-S., Lee, J.-E., Shin, J., & Kim, M. 2014, Systematic effects on the Genus topology of the Large-scale Structure of the Universe, *ApJS*, 212, 22
- Kim, J., & Park, C. 2006, A new halo-finding method for N-body simulations, *ApJ*, 639, 600
- Kim, J., Park, C., Gott, J. R., & Dubinski, J. 2009, The Horizon Run N-body simulation: baryon acoustic oscillations and topology of Large-scale Structure of the Universe, *ApJ*, 701, 1547
- Kim, J., Park, C., & Choi, Y.-Y. 2008, A subhalo-galaxy correspondence model of galaxy biasing, *ApJ*, 683, 123
- Kim, J., Park, C., Rossi, G., Lee, S.M., & Gott, J.R. 2011, The new Horizon Run cosmological N-body simulations, *Journal of Korean Astronomical Society*, 44, 217
- Klypin, A., Prada, F., Yepes, G., Heß, S., & Gottlöber, S. 2015, Halo abundance matching: accuracy and conditions for numerical convergence, *MNRAS*, 447, 3693
- Kravtsov, A., Vikhlinin, A., & Meshcheryakov, A. 2014, Stellar mass – halo mass relation and star formation efficiency in high-mass halos, arXiv:1401.7329v1
- L'Huillier, B., Park, C., & Kim, J. 2014, Effects of the initial conditions on cosmological N-body simulations, *New Astronomy*, 30, 79
- Lacey, C. & Cole, S. 1993, Merger rates in hierarchical models of galaxy formation, *MNRAS*, 262, 627
- Lee, J. et al. 2014, Sussing merger trees: the impact of halo merger trees on galaxy properties in a semi-analytic model, *MNRAS*, 445, 4197
- Lewis, A., Challinor, A., & Lasenby A. 2000, Efficient computation of cosmic microwave background anisotropies in closed Friedmann-Robertson-Walker models, *ApJ*, 538, 473
- Lukić, Z., Heitmann, K., Habib, S., Bashinsky, S., & Ricker, P. 2007, The halo mass function: high-redshift evolution and universality, *ApJ*, 671, 1160
- Manera, M., Sheth, R. K., & Scoccimarro, R. 2010, Large-scale bias and the inaccuracy of the peak-background split, *MNRAS*, 402, 589
- Monaghan, J. J. 1992, Smoothed particle hydrodynamics, *Annual Reviews of Astronomy and Astrophysics*, 30, 54
- Muldrew, S. I., Pearce, F. R., & Power, C. 2011, The accuracy of subhalo detection, *MNRAS*, 410, 2617
- Onions, et al. 2012, Subhaloes going Notts: the subhalo-finder comparison project, *MNRAS*, 423, 1200
- Park, C., Choi, Y.-Y., Kim, J., Gott, J. R., Kim, S. S., & Kim, K.-S. 2012, The challenge of the largest structures in the universe to cosmology, *ApJ*, 759L, 7
- Reddick, R. M., Wechsler, R. H., Tinker, J. L., & Behroozi, P. S. 2013, The connection between galaxies and dark matter structures in the Local Universe, *ApJ*, 771, 30
- Rigopoulos, G. & Valkenburg, W. 2015, On the accuracy of N-body simulations at very large scales, *MNRAS*, 446, 677
- Rodríguez-Puebla, A., Avila-Reese, V., & Drory, N. 2013, The galaxy-halo/subhalo connection: mass relations and implications for some satellite occupational distributions, *ApJ*, 767, 92
- Rossi, G., Sheth, R. K. & Tormen, G. 2011, Modelling the shapes of the largest gravitationally bound objects, *MN-*

- RAS, 416, 248
- Sheth, R. K. & Tormen, G. 1999, Large-scale bias and the peak background split, MNRAS, 308, 119
- Smith, R. E., Peacock, J. A., Jenkins, A., White, S. D. M., Frenk, C. S., Pearce, F. R., Thomas, P. A., Efstathiou, G., & Couchman, H. M. P. 2003, Stable clustering, the halo model and non-linear cosmological power spectra, MNRAS, 341, 1311
- Speare, R., Gott, J. R., Kim, J., & Park, C. 2015, Horizon Run 3: topology as a standard ruler, ApJ, 799, 176
- Tatekawa, T. & Mizuno, S. 2007, Transients from initial conditions based on Lagrangian perturbation theory in N-body simulations, Journal of Cosmology and Astroparticle Physics, 12, 014
- Tinker, J., Kravtsov, A. V., Klypin, A., Abazajian, K., Warren, M., Yepes, G., Gottlöber, S. & Holz, D. E. 2008, Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality
- Trujillo-Gomez, S., Klypin, A., Primack, J., & Romanowsky, A. J. 2011, Galaxies in CDM with halo abundance matching: luminosity-velocity relation, baryonic mass-velocity relation, velocity function, and clustering, ApJ, 742, 16
- Villalobos, Á., De Lucia, G., Weinmann, S. M., Borgani, S., & Murante, G. 2013, An improved prescription for merger time-scales from controlled simulations, MNRAS, 433, L49
- Wang, Y. O., Lin, W. P., Kang, X., Dutton, A., Yu, Y., & Macciò, A. V. 2014, Satellite alignment. I. distribution of substructures and their dependence on assembly history from N-body simulations, ApJ, 786, 8
- Warren, M. S., Abazajian, K., Holz, D. E., & Teodoro, L. 2006, Precision determination of the mass function of dark matter halos, ApJ, 646, 881
- Watson, W. A., Iliev, I. T., D'Aloisio, A., Knebe, A., Shapiro, P. R., & Yepes, G. 2013, The halo mass function through the cosmic ages, MNRAS, 433, 1230
- Weinberg, D. H., Colombi, S., Davé, R., & Katz, N. 2008, Baryon dynamics, dark matter substructure, and galaxies, ApJ, 678, 6
- Zhao, H. 2004, Dynamical friction for dark halo satellites: effects of tidal mass loss and growing host potential, MNRAS, 351, 891
- Zheng, Z., et al. 2005, Theoretical models of the halo occupation distribution: separating central and satellite galaxies, ApJ, 633, 791
- Zheng, Z., Zehavi, I., Eisenstein, D. J., Weinberg, D. H., & Jing, Y. P. 2009, Halo occupation distribution modeling of clustering of luminous red galaxies, ApJ, 707, 554